Discussion Papers
Department of Economics
University of Copenhagen

No. 17-04

Choosing a Public-Spirited Leader. An experimental investigation of political selection

Thomas Markussen

Jean-Robert Tyran

Øster Farimagsgade 5, Building 26, DK-1353 Copenhagen K., Denmark
Tel.: +45 35 32 30 01 – Fax: +45 35 32 30 00
http://www.econ.ku.dk

ISSN: 1601-2461 (E)

Choosing a Public-Spirited Leader

An experimental investigation of political selection

Thomas Markussen and Jean-Robert Tyran*

March 2nd, 2017

Abstract: In this experiment, voters select a leader who can either act in the public interest, i.e. make efficient and equitable policy choices, or act in a corrupt way, i.e. use public funds for private gain. Voters can observe candidates' pro-social behavior and their score in a cognitive ability test prior to the election, and this fact is known to candidates. Therefore, self-interested candidates have incentives to act in a pro-social manner, i.e. to pretend to be public-spirited leaders. We find that both truly pro-social and egoistic leaders co-exist, but that political selection is ineffective in choosing public-spirited leaders. The main reason is that egoistic candidates strategically pretend to be pro-social to increase their chances of winning the election.

Keywords: Political selection, pro-social behavior, social dilemma, corruption, voting

JEL codes: C92, C91, D03, D72, H41

---

* Markussen: Department of Economics, University of Copenhagen, Øster Farimagsgade 5, 1353 København K, Denmark. Tyran: Department of Economics and VCEE, University of Vienna, Oskar-Morgenstern-Platz 1, 1090 Vienna, Austria and Department of Economics, University of Copenhagen, Denmark and CEPR London. Tyran thanks the Austrian Science Fund under project number I2027-G16 and Markussen the Danish Council for Independent Research (FSE) for generous financial support. We thank Marco Piovesan for his support in running the experiment and Rebecca Morton and seminar participants at the University of Copenhagen and the University of Potsdam for helpful comments.

Acting is all about honesty. If you can fake that, you've got it made.

George Burns

1. Introduction

How do voters select political leaders? Under what conditions does this selection process result in "good" politicians being in charge? These questions are of obvious practical relevance for the functioning of democracy but they have received the attention they deserve by academic (political) economists only in recent years (e.g. Besley 2006, Caselli and Morelli 2004, Messner and Polborn 2004, Diermeier, Keane and Merlo 2005, Dal Bo, Dal Bo and Di Tella 2007, Matozzi and Merlo 2007, Acemoglu, Egorov and Sonin 2011, Ferraz and Finan 2009, Galasso and Nannicini 2011, Brollo et al. 2013). Studies of political selection typically investigate the determinants of the "quality" of elected politicians where quality is measured, for example, in terms of competence or honesty. Selection may take place either at the stage of entry into the political arena or the stage of actual selection for office, for example through an election (see Besley 2005). While a number of empirical studies of political selection exist, only few have used experimental methods which allow researchers to precisely pin down determinants in stylized environments (e.g. Dasgupta and Williams 2002, Houser, Ludwig and Stratmann 2009, Hamman, Weber and Woon 2011, Corazzini et al. 2014, Galeotti and Zizzo 2015; see section 2 for a discussion).

This paper focuses on selection of public-spirited versus corrupt leaders at the election stage. In our setting, all voters prefer public-spirited leaders and all politicians have incentives to be corrupt: leaders are in control of public funds and can use those either in the public interest - by providing a public good which is both efficient and fair - or for their exclusive private gain. It is a fairly common view that all candidates for political office are corrupt "rascals". However, this view is not plausible in theory (Bernheim and Kartik 2016), nor is it supported in experimental research (e.g. Drazen and Ozbay 2016). We therefore assume (and we confirm this assumption in our results) that voters indeed have the choice between candidates who are public-spirited and purely self-interested. The voters' problem is that they cannot directly observe the candidates' "types". Instead, candidates can signal their type to voters through their reputations, i.e. their observable behavior before the election. This creates an incentive for self-interested candidates to pretend to be public-spirited in the pre-election period in order to improve their electoral prospects. Strategic pretending, in turn, hampers voters' ability to select truly public-spirited leaders.

We design an experiment to investigate whether such strategic pretending takes place, and how it affects the quality of political selection. The experimental method is well suited for this type of investigation. In the field, it is hard to adequately measure the pro-social behavior of political candidates and to determine whether, once elected, their policies serve the public interest or not. In particular, it is difficult to know the counterfactual policies that candidates *not* elected would have chosen, had they been elected. Without this information, it is hard to judge the quality of political selection, i.e. whether voters in fact succeed in choosing the most public-spirited leaders. In contrast, our experiment measures social behavior that is unobserved by voters as a proxy for pro-social orientation. In our setting it is easy to determine whether policies serve the public interest, and we learn the policy choices of both winners and losers in an election. Our experimental design also allows us to gauge whether candidates engage in strategic pretending in the pre-election period, and how this behavior affects the political selection process.

In essence, our experimental design consists of a vote which takes place between two public goods (PG) games. First, subjects play a standard PG game in which each group member is in control of his or her endowment and can allocate it to a private good or the PG. Choices in this game are observed by voters, and all participants are aware of this fact from the beginning of the game. Voters then elect one subject as the leader who controls all endowments in the group, and decides how much of the total endowments to allocate to the PG, and how much to keep for himself in the second PG (the "centralized" PG game). Hence, strictly self-interested participants may have an incentive to strategically contribute in the standard PG game to improve their chances of winning the election.

Candidates are required to commit to their behavior as leaders before knowing whether they in fact won the election (that is, we use the "strategy method"). The advantage of this procedure is that the policies of winners as well as losers in the election are known to the experimenter. This allows us to analyze whether voters are successful in picking the most public-spirited candidates. In control treatments, leaders are randomly appointed. This benchmark allows us to investigate whether efficiency is higher with elected leaders than with randomly appointed ones. In addition, we let subjects play several rounds of the public goods game knowing that they are unobserved by others, and not knowing that a voting game follows, before the game described above begins. This unobserved part provides information about participants' pro-social behavior that is free of strategic incentives as a benchmark.

Our main findings are that political selection is potentially important, in the sense that there is high and partly predictable variation in the behavior of leaders. The modal leader is corrupt and simply pockets the entire endowment, but about 18 percent of leaders are public-spirited in that they allocate the entire endowment to the public good. It therefore matters a great deal who gets elected. Voters are to a large extent able to use the information they are provided rationally. They tend to base their voting decisions on the observed contributions to public good (but not on a score for cognitive reflection, which is also provided). Yet, the political selection process in the experiment turns out to be ineffective. We find that the chosen leaders are not significantly more pro-social than the non-chosen, or than leaders selected at random. We argue that the key explanation for this result is strategic contribution behavior by self-interested candidates in the pre-election period. Strategic pretending hampers voters' ability to select the most public-spirited types.

To cleanly isolate issues of political selection and strategic behavior, we control for a number of aspects of politics which are likely to matter much in the wild, such as incumbency and re-election incentives. The literature section below references papers that did address these issues, and we discuss potential extensions of our set-up in the conclusion.

The paper is structured as follows. Section 2 discusses related literature and section 3 presents the experimental design. Section 4 derives theoretical predictions and section 5 presents results. Section 6 concludes.

2. Related studies

This study is related to a group of theoretical papers analyzing signaling games between politicians and voters (e.g. Austen-Smith and Banks 1993, Banks and Sundaram 1993, Coate and Morris 1995, Fearon 1999, Besley 2006, Besley and Smart 2007, Kartik and McAfee 2007, Markussen and Tyran 2010). With the exception of Kartik and McAfee and Markussen and Tyran, however, these studies mostly focus on signaling by a single, *incumbent* office holder. In contrast, we focus on signaling by several, non-incumbent candidates.

Turning to experimental papers, a group of related studies investigate the effects of political campaigns. Corazzini et al. (2014) study a set-up where an elected leader distributes a budget between him- or herself and an electorate. This feature is similar to our design. Before the election, candidates are allowed to make non-binding "campaign promises" stating how large a share of the budget they will allocate to voters. In one control treatment, leaders are randomly appointed. In

another, the possibility of making campaign statements is shut down. The main differences from our experiment are, first, that signaling is free ("cheap talk"), whereas in our experiment it is costly. Second, signals in Corazzini et al. take the form of explicit promises. This may be important if leader behavior is affected by "guilt aversion", as argued by the authors. The breaking of an explicit promise is likely to evoke stronger feelings of guilt than the act of simply behaving more egoistically after the election than before, as candidates may do in our experiment. The results in Corazzini et al. show that elected leaders act more pro-socially than leaders who are appointed randomly. This effect is only present when candidates are allowed to campaign. This result differs from ours in the sense that elected leaders in our experiment do not turn out to act more benevolently than randomly appointed leaders. Due to the differences in the experimental design, however, results in Corazzini et al. and the present paper are not directly comparable.

Houser, Ludwig and Stratmann (HLS 2009) investigate how deceptive campaign advertising affects voter turnout and efficiency of electoral outcomes. In line with our findings, they report that the presence of deceptive information strongly reduces the efficiency of the democratic process. A main difference from our experiment is that candidates in HLS are "virtual" players. No real subjects are assigned to the role of candidates. Consequently, there are no policy choices and no strategic behavior by candidates. The computer simply determines randomly whether a "candidate" sends a true or a false signal and payoffs follow automatically from the election result. In contrast, strategic behavior and policy choices by "real" candidates are key features of our study. Houser and Stratmann (2008) and Houser, Morton and Stratmann (2011) investigate effects of offering costly vs. free information to voters. These studies differ from ours in that campaign information is always truthful. Therefore, in contrast with our set-up, a low quality candidate cannot pretend to be of high quality.

Dasgupta and Williams (2002) investigate whether opinion polls are an effective means of transmitting information about candidate quality from an informed minority to the uninformed majority. Their study shares with ours the feature that candidates may engage in costly signaling to convince voters that they are of high quality. However, in contrast with our study, candidate quality is exogenous, and candidates do not make any policy choices after the election.

Hamman, Weber and Woon (2011) is a close match in that they also study a public goods game where the authority to decide on allocations to the public good is in some treatments delegated to one individual who is chosen in an election. Their design differs from ours by featuring

repeated elections (and therefore re-election incentives) and by presenting voters with information on contributions to the public good in an initial, standard public goods game, where incentives for strategic contributions were absent, because subjects were not aware that phases with electoral competition were coming up later in the experiment. In contrast to our study, they do not include control treatments with (randomly) appointed rather than elected leaders. Results show that elected leaders act quite pro-socially and that efficiency is much higher with elected leaders than in a standard public goods game.

Galeotti and Zizzo (2015) use a real-effort experiment to study voter preferences for competence and honesty, respectively, of candidates for political office. Voters' payoff depends on a) the performance ("competence") of an elected official in a real-effort task, and b) willingness of the official to report honestly how much he or she produced in the task (underreporting generates rents for the official). Voters are informed about the competence and honesty of candidates in a previous phase of the experiment. Results show that voters prioritize honesty over competence, to the point of frequently voting for the most honest candidate even when that candidate is not expected to deliver the highest monetary payoff to the voter.

A bunch of experimental papers have investigated whether leaders become more pro-social when elected or volunteering rather than when randomly chosen in various contexts. The results are mixed. Levy et al. (2011), Brandts, Cooper and Weber (2014) and Drazen and Ozbay (2016) find such effects, Bolle and Vogel (2011) do not.

Finally, a number of experimental studies have examined whether groups are willing and able to implement efficiency-enhancing institutions in social dilemma situations (e.g. Walker et. al. 2000, Tyran and Feld 2006, Okada, Kosfeld and Riedl 2009, Dal Bo, Foster and Putterman 2011, Sutter, Haigner and Kocher 2010, Kamei, Putterman and Tyran 2010, Markussen, Putterman and Tyran 2014, Markussen, Reuben, and Tyran 2014). We also consider voting as a potential means to solve collective action problems, but we focus on endogenous *leaders*, rather than endogenous institutions.

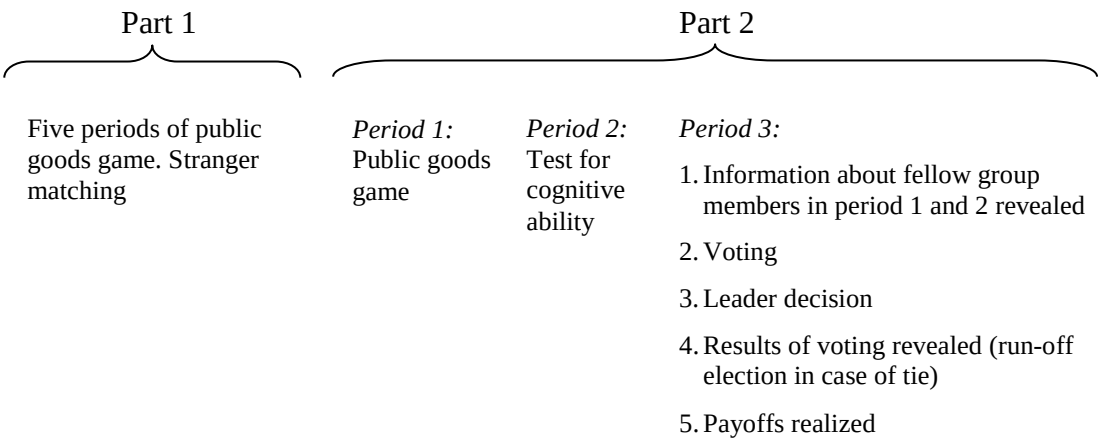
3. Experimental design

In total, 234 freshmen students of economics (about one and a half months into their studies) participated in 11 sessions with an average of 21.3 participants in each (from minimum of 16 to a maximum of 28) at the University of Copenhagen. The experiment lasted about one and a half

hours, and participants earned on average 78 Danish kroner (about 16 dollars). We paid no show-up fee. The experiment was programmed and conducted using the software z-Tree (Fischbacher 2007). Instructions used neutral language. For example, public goods are referred to as "group accounts", and the person referred to here as the "leader" is called the "decision maker".

Figure 1 provides an overview of timing in the experiment, which has two parts. The first part consists of five periods of a standard, linear public goods game. It serves to familiarize participants with the logic of voluntary provision of public goods and the free-riding incentives that come with it. Part 2 is the main part in which reputation can be built strategically, political selection of the leader takes place, and the leader chooses the allocation on behalf of the group. In part 1, participants earn one Danish krone for every five experimental points they collect, while in part 2 they earn one krone for each experimental point. The exchange rate is increased in part 2 to concentrate financial incentives in the part of the experiment which is our main focus of attention. However, studies of the effects of changes in monetary stakes on behavior in laboratory experiments tend to find weak effects (e.g. Camerer and Hogarth 1999, Carpenter, Verhoogen and Burks 2005).

Figure 1: Timing in main treatment



In part 1, participants are divided into groups of three. In each period, groups are randomly reshuffled ("stranger matching"). Participants receive an endowment of 20 points in each period. They decide how much of the endowment to allocate to a public good or to keep for themselves. The payoff function for participant i in group g is

$$Y_i = 20 - C_i + 0.5\sum_{j \in g} C_j$$

where C_i is participant i's contribution to the public good. It is easy to verify that in each round it is a dominant strategy for an agent maximizing the expected value of Y_i to contribute nothing to the public good, while the social optimum (i.e. the maximum sum of earnings) is reached if all participants contribute their entire endowment to the public good.

Instructions for part 2 are only handed out after the completion of part 1. In the first period of part 2, participants play another round of a linear public goods game with the same structure as in part 1 but knowing about all the elements and the information provided about them that follow in part 2. Between period 5 of part 1 and period 1 of part 2, groups are again randomly reshuffled. In period 2 of part 2, participants take a test for cognitive ability involving 15 logical puzzles, sampled from a set of tasks typically used in IQ tests (taken from Kamei et al. 2015). They earn one point per correct answer (average is 6.5, see figure B2 in Appendix B for the distribution).

At the beginning of period 3 of part 2, groups are again randomly reshuffled. Leaders are elected and make their policy decisions. The public goods game of period 3 is similar to that in period 1, but instead of each participant in a group receiving an endowment of 20 points, the leader receives all 60 points, while the two others receive nothing. The leader decides how much to allocate to the public good, and how much to keep for himself. Each group member, including the leader, receives 0.5 times the amount allocated to the public good. Hence, the leader earns between 30 and 60 points, while other group members earn between 0 and 30 points. This means that the leader faces a social dilemma: Allocating the full endowment to the public good maximizes efficiency as well as equity, but the leader's own earnings are maximized if he keeps the entire endowment for himself.

In the main treatment, the leader is chosen by voting. All group members are obliged to vote, and it is not possible to vote for oneself. Before votes are cast, participants learn how much their fellow group members contributed to the public good in period 1 of part 2, and how many puzzles they solved correctly in the test in period 2.[1]

[1] The exact wording used in the table presenting information about candidates is "allocation to the group account in stage 1" and "number of correct answers in quiz in stage 2" (see decision screens in Appendix D).

This is the stage where political selection potentially plays a role. If behavior in period 1 and 2 is correlated with leader behavior in period 3, the information could enable voters to choose the most qualified leaders. On the other hand, since participants are informed about these rules *before* they make their contribution decision in period 1, this feature of the experiment also creates an incentive for strategic behavior – some players might choose a high contribution in period 1 to increase their chances of winning the election and to be in control of the endowments in period 3. By comparing a participant's behavior in period 1 of part 2 with his or her behavior as leader, we can test if the leader was in fact motivated by strategic considerations.

In case the election is tied (i.e. each group member receives one vote), a run-off election is held where one group member is randomly excluded from running (but not from voting). In the run-off election, voting for oneself is allowed. All group members are again presented with the period 1 and 2 information about the candidates before they cast their vote. The advantage of these voting rules are, first, that all experimental subjects are eligible, i.e. are in the role of a "candidate". This means that the amount of data on the behavior of candidates and leaders is maximized. Second, the rules of the run-off election ensure that a winner emerges.[2]

Our measure of cognitive ability is a proxy for candidate "competence". Since the cognitive ability score does not affect the size of the endowment that leaders get to distribute, we may expect cognitive score to be a less important predictor of leader performance than behavior in the stage 1 public goods game (i.e. our measure of "public-spiritedness"). Still, information about candidates' cognitive ability may be relevant for at least two reasons. First, candidates with high ability are more likely to understand the rules. There may be some comfort to voters in knowing that their leader will at least make an informed decision. Second, leaders with high ability scores may choose higher contributions to the public good due to "inequality aversion" (cf. Fehr and Schmidt 1999). In particular, a candidate with a high test score may feel that he or she has already received a significant, non-monetary benefit by being presented to the other group members as a smart person. If the candidate is averse to advantageous inequality of non-monetary as well as monetary benefits,

---

[2] The fact that candidates in the run-off election are allowed to vote for themselves allows us to still present the procedure as an "election". The alternative would be to allow only the excluded candidate to vote, but in that case the term "election" might be confusing because there would be only one voter. On the other hand, if candidates were allowed to vote for themselves in the first round of the election, it is likely that most candidates would do so, and that we would therefore learn little about the effects of information about candidates on electoral success.

he or she may feel that it is fair to compensate the other group members by giving them a high share of the stage 3 endowment.

On the other hand, it is also possible that the ability score *negatively* relates to leader contributions. In particular, high scores may trigger a feeling of entitlement in participants with high scores. For example, Hoffman et. al. (1994) found that when the role of dictator in a dictator game is allocated based on score in a trivia quiz, dictator allocations to the receiving player go down, compared with the standard situation where roles are randomly allocated.

In addition to the possibility that test scores affect the pro-social behavior of leaders in one direction or the other, there is also a methodological advantage attached to presenting voters with this information: if players were only presented with data on public goods contributions in period 1, it is likely that they would vote for high contributors even if they did not interpret high contributions as evidence of pro-social preferences, simply because high contributions would be "focal". However, when we add the attention-catching information on cognitive ability scores, it is more likely that players will only vote for high period 1 contributors if they actually do interpret high contributions as evidence of altruism. Therefore, we provide a more stringent test of our hypothesis of voters' ability to select candidates according to relevant criteria by adding information on cognitive ability.

Before the result of the election is revealed, all participants indicate how much they will contribute to the public good given that they are chosen to be the leader. The winner of the election is then announced, and his or her policy is implemented. This sequence ensures that we know the policy choices of both winners and losers of the election, and are able to determine whether voters succeed in picking the most public-spirited candidates.

***Table 1*:** Treatments (number of observations in parentheses)

		Information about stage 1 and 2	
		Yes	No
Vote	Yes	Main treatment (n = 126)	Treatment 2 (n = 33)
	No	Treatment 3 (n = 33)	Treatment 4 (n = 42)

Political selection is made possible in the main treatment by the combination of elections *and* the information about candidates given to voters. To isolate the effect of *combining* these two elements, we introduce three types of control treatments: one where an election is held, but no information about candidates is provided, one where information is provided, but no election is held, and one where no election is held, and no information is provided. Table 1 provides an overview of the four treatments. In each experimental session, all four treatments are used. Participants are informed that they will be randomly allocated to treatments, and that the probability of being in the main treatment is 50 percent, while being in each of the other three treatments is equally likely.[3] Participants only learn which treatment they are in at the beginning of period 3. At the beginning of the experiment, and again at the beginning of period 3 in part 2, participants answer a set of questions testing their understanding of the rules of the game.

4. Theoretical predictions

Consider a slightly simplified version of the game played in part 2 of the experiment. In particular, assume that a population of N agents is divided into groups of 3. The groups play a two-period public goods game with the same parameters as in the experiment (endowment per person, per period = 20; marginal per capita return = .5). Groups are re-shuffled between the two periods. In the second period, all endowments are given to one player, the leader, who decides how much to contribute to the public good. The leader is appointed through voting. All group members are candidates in the election, and voting for one-self is not allowed. With a probability of .5, group members in period 2 are informed about each other's contributions to the public good in period 1. This is equivalent to the probability of being in the main treatment in the experiment.

Assume that voters always vote "sincerely" in this game, i.e. vote for the candidate they genuinely prefer to see winning. Assume, furthermore, that two types of agents exist: First, *pro-social types* always choose the maximum contribution to the public good, both in period 1 and period 2. *Egoistic types*, on the other hand, simply maximize their expected payoff. They therefore contribute nothing as leaders in period 2. In period 1, they only contribute if doing so maximizes total expected payoff. An egoist may wish to contribute in order to improve his chances of winning

[3] Actual numbers deviate from these percentages due variation in number of subjects across sessions.

the election in period 2. We refer to an egoist who chooses a high contribution in period 1 as a *strategic contributor*. Since all pro-social agents choose the maximum contribution in period 1, voters always vote for a high period 1 contributor, if they get the chance. Denote the share of pro-social types in the population by θ. The value of θ is common knowledge.[4]

Denote the contribution of agent *i* to the public good in period 1 by C_i. Denote the agent's earnings in period 1 by x_{i1}, his earnings in period 2 by x_{i2} and his total earnings by x_i. Now, an egoist chooses $C_i = 20$ if and only if

$$\begin{aligned} &E(x_i \mid C_i = 20) \geq E(x_i \mid C_i = 0) \\ &\Rightarrow E(x_{i1} \mid C_i = 0) - E(x_{i1} \mid C_i = 20) \leq E(x_{i2} \mid C_i = 20) - E(x_{i2} \mid C_i = 0) \end{aligned} \tag{1}$$

where *E* denotes the expectations operator. In words, the expected loss from a high contribution in period 1 must be lower than the expected gain in period 2. Note that the loss associated with a high contribution is fixed at exactly .5 x 20 = 10 points.

Expected earnings in period 2 depend on the probability of being elected leader and on the expected earnings if one is not elected. An egoist earns 60 points if elected leader. If he is not elected, he will earn zero if the leader in his group is an egoist and 30 points if the leader is pro-social. Denote by *P* the probability that the leader in *i*'s group is pro-social, given that *i* is not elected himself. Define the dummy variable D_i to be one if player *i* is the leader and zero otherwise. We then have:

$$\begin{aligned} E(x_{i2} \mid C_i) &= prob(D_i = 1 \mid C_i)60 + (1 - prob(D_i = 1 \mid C_i))30P \\ &= prob(D_i = 1 \mid C_i)(60 - 30P) + 30P \end{aligned} \tag{2}$$

Because information about period 1 actions is only revealed before the election in period 2 with probability 1/2, the gain in period 2 earnings associated with a high period 1 contribution is:

$$E(x_{i2} \mid C_i = 20) - E(x_{i2} \mid C_i = 0) = \frac{1}{2}\big(prob(D_i = 1 \mid C_i = 20) - prob(D_i = 1 \mid C_i = 0)\big)(60 - 30P) \tag{3}$$

Now, the probability that *i* is elected leader depends on the number of high period 1 contributors in his period 2 group. Denote the number of high contributors in the population by *H*

---

[4] Markussen and Tyran (2009) present a related but more general model on strategic behavior and political selection.

and the number of high contributors in *i*'s group, apart from himself, by H_g. Voters always prefer to vote for a high contributor. Therefore, if *i* chooses $C_i = 20$, he will be elected with probability 1 if $H_g = 0$, with probability 1/2 if $H_g = 1$, and probability 1/3 if $H_g = 2$. If he chooses $C_i = 0$ (or any other contribution lower than 20), he will be elected with probability 1/3 if $H_g = 0$. If $H_g > 0$ he will not be elected. Therefore,

$$\begin{aligned} prob(D_i = 1 \mid C_i = 20) - prob(D_i = 1 \mid C_i = 0) &= \frac{2}{3} prob(H_g = 0) + \frac{1}{2} prob(H_g = 1) + \frac{1}{3} prob(H_g = 2) \\ &= \frac{2}{3} - \frac{H}{3(N-1)} \end{aligned} \tag{4}$$

(Proof in Appendix A). This means that the increase in the probability of being elected associated with choosing $C_i = 20$ instead of $C_i = 0$ depends on the number of other high contributors. It decreases from 2/3 when no-one else contributes to 1/3 when everyone else contributes.

Now consider *P*, the probability that the leader in *i*'s group is pro-social, given that *i* is not elected himself. Players who are pro-social choose a high contribution in period 1 for sure, but not all players who choose full contributions are pro-social. Out of the total pool of high contributors, *H*, only (approximately) $\theta(N-1)$ are in fact pro-social. *P* can therefore be expressed as the probability that there is at least one high contributor in a group multiplied by the probability that a given high contributor is pro-social:

$$\begin{aligned} P &= (1 - prob(H_g = 0)) \frac{\theta(N-1)}{H} = \left(1 - \left(1 + \left(\frac{H}{N-1} \right)^2 - \frac{2H}{N-1} \right) \right) \frac{\theta(N-1)}{H} \\ &= \theta \left(2 - \frac{H}{N-1} \right) \end{aligned} \tag{5}$$

(See the proof of (4) in Appendix A for a proof of the expression for $prob(H_g = 0)$). Hence, *P* depends positively on θ, but negatively on *H*. Given θ, a higher number of high contributors in period 1 means that more egoists strategically pretend to be pro-social. This pretending makes it more difficult to choose a truly pro-social candidate in the election in period 2. Hence, the total effect of *H* on the period 2 benefit from choosing a high contribution in period 1 is ambiguous: on the one hand, a higher *H* means that the effect of a high contribution on the probability of being

elected leader is weaker. On the other hand, a higher *H* also increases the expected cost of not being elected (for a given θ).

Inserting (4) and (5) in (1) and (3) and re-writing, we get the following condition for egoist *i* to choose $C_i = 20$:

$$\frac{H}{(N-1)} \leq \frac{2\theta - 1 + \sqrt{1-2\theta}}{\theta} \equiv t \tag{6}$$

(Proof in Appendix A). Expression (6) shows that it is only optimal for an egoist to choose $C_i = 20$ when the share of other players choosing a high contribution is below a certain threshold, which we denote by *t*. Hence, the negative effect of *H* on the probability of getting elected dominates the positive effect on the cost of not being elected. When $t \leq \theta$, no egoists will contribute (i.e. a "separating equilibrium" prevails). When $t \geq 1$, all players contribute (i.e. a "pooling" equilibrium prevails). When $\theta < t < 1$, some but not all egoists contribute. In particular, if a share of egoists equal to ($t - \theta$) / (1 – θ) contribute, and the complementary share does not, then no-one has an incentive to deviate. This equilibrium can be implemented, for example, if all egoists play a mixed strategy where they choose a high contribution with probability ($t - \theta$) / (1 – θ). When θ > ½, the threshold *t* does not exist. In those cases, it does not pay for egoists to choose a high contribution.

Now consider how strategic pretending affects the effectiveness of voting as a means to selecting the best leaders. We can compare the situation with information and strategic behavior, which we have modeled to two benchmarks: first, the situation with no information about social preferences of candidates, where selection of leaders is random, as in the control treatments in the experiment. In this case, the probability that an elected leader is pro-social, denoted $\tilde{P}_{random}$, is equal to θ. Second, consider the situation with information, but no strategic behavior. Imagine that some technology is available which allows voters to distinguish perfectly between pro-social agents and egoists. Then, all groups with at least one pro-social member will succeed in choosing a public-spirited leader. The probability that a such a leader is selected is therefore equal to the probability that at least one group member is pro-social, which equals $\tilde{P}_{selection} = 1 - (1-\theta)^3$.

In the situation with both information and strategic behavior, the probability that a group elects a pro-social leader equals the probability that the group has at least one high period 1

contributor, multiplied by the probability that a high contributor is pro-social. This gives the following expression for $\tilde{P}$:

$$\tilde{P}_{strategic} = \begin{cases} 1-(1-\theta)^3, & t < \theta \\ \left(1-(1-t)^3\right)\dfrac{\theta}{t}, & \theta \le t \le 1 \\ \theta \quad , & t > 1 \end{cases} \qquad (1)$$

To interpret this expression, remember that when $t < \theta$, there is no strategic behavior. When $t \ge 1$, all egoists strategically contribute, and voters receive no useful information. We have that $\tilde{P}_{selection} \ge \tilde{P}_{imitation} \ge \tilde{P}_{random}$. (For the first inequality, the proof for $\theta \le t \le 1$ follows from the fact that

$$\frac{d\left((1-(1-x)^3)/x\right)}{dx} = 2x - 3 < 0 \quad \text{for } 0 \le x \le 1).$$

Figure 2: Appointment technology and the probability of electing a pro-social leader

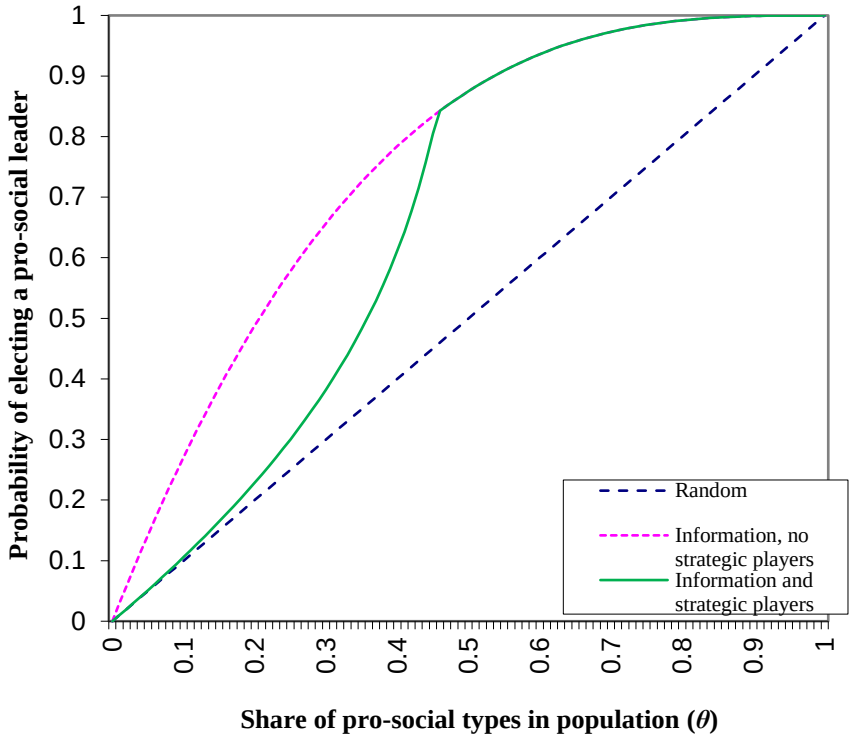


Figure 2 shows $\tilde{P}_{selection}$, $\tilde{P}_{imitation}$ and $\tilde{P}_{random}$ as functions of θ. When there is no strategic behavior, voting combined with information about social preferences leads to a substantial improvement in the probability of electing a good leader, as compared to the outcomes from random appointment. This is true even with a low share of pro-social agents in the population. However, when strategic behavior by egoists is admitted, these benefits are almost completely eliminated for low values of θ. On the other hand, for high values of θ, the cost of strategic behavior is small or non-existent.

In other words, it is an empirical question how much damage strategic behavior will cause by undermining political selection. The experiment gives an empirical answer to this question.

The model predicts that there will be a mixed equilibrium among egoists, with some contributing in period 1 and others not, if the share of pro-social agents is between 0 and 45 percent. Figure 2 shows that strategic behavior will severely damage the ability of voters to select public-spirited leaders if the share of pro-social agents falls below 40 percent.

5. Results

a) Leader behavior – is political selection important?

Since our main focus is whether groups are able to select pro-social leaders, we start by presenting the distribution of leaders' contributions, and then analyze the determinants of these contributions.

Figure 3: Leaders' contributions (n = 234)

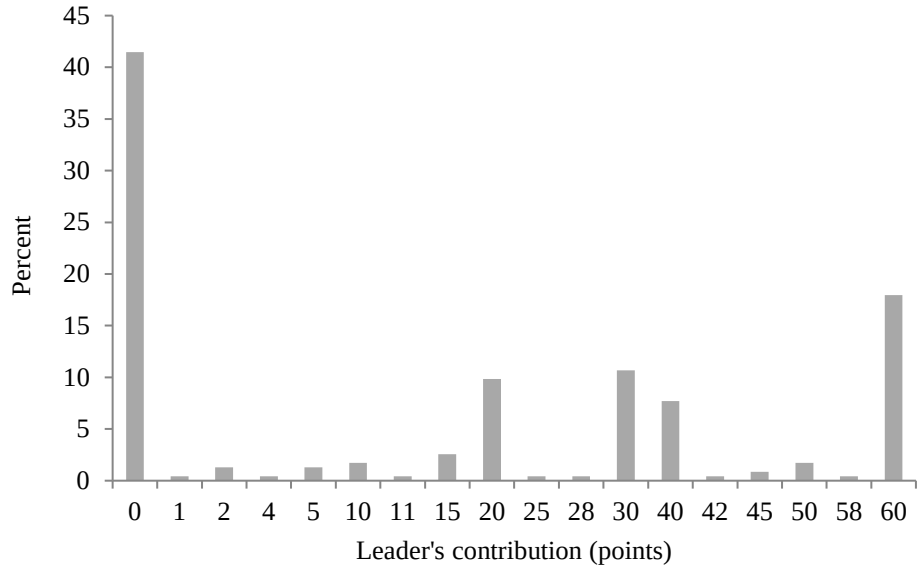


Figure 3 shows that a majority of potential leaders was not fully corrupt, but only few were fully public-spirted either. More precisely, the figure shows the distribution of contribution choices of all participants conditional on being elected leaders, i.e. their choices of what they would do if they were chosen to be leader. Standard economic theory predicts that all participants allocate zero points to the public good. Zero is indeed the modal strategy (42 percent), but a clear majority of leaders (58 percent) allocate a positive amount to the public good. On average, participants allocate 36 percent (22 points) of the 60 point endowment to the public good. 18 percent of leaders allocate the entire endowment to the public good. The maximum contribution is the secondary mode. If all groups had perfect foresight and elected their member with the highest leader contribution, leaders would on average have contributed 68 percent of the endowment, or almost twice as much as the average participant. Hence, there is a potential for political selection to play an important role. In what follows we investigate whether this potential was realized.

Table 2 shows that leader contributions can be predicted on the basis of contributions to the PG immediately preceding the vote, but less so on the basis of the cognitive ability test (see Figures B1 and B2 in the appendix for the respective distributions).

Table 2: Determinants of leader behavior

	Dependent variable: Contribution as leader (points)					
	All participants			Participants in main treatment		
Contrib. to PG in part 2, period 1	0.692 (3.78)***		0.713 (3.89)***	0.655 (2.65)***		0.703 (2.86)***
Correct answers in test		-0.584 (1.05)	-0.756 (1.39)		-1.173 (1.51)	-1.405 (1.84)*
Constant	15.788 (7.40)***	25.421 (6.47)***	20.526 (5.11)***	16.981 (5.95)***	29.892 (5.57)***	25.546 (4.70)***
Observations	234	234	234	126	126	126
R-squared	0.06	0.00	0.07	0.05	0.02	0.08

Notes: OLS regressions. Absolute value of *t*-statistics in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%

For the full group of participants in the experiment, contributions to the public good in period 1 of part 2 is a highly significant predictor of leader behavior, while there is no significant effect of performance in the cognitive ability test. Among participants in the main treatment, the effect of period 1 contributions remains highly significant. However, test score is now also significant (at the 10 percent level), when period 1 contributions are controlled for. Higher test scores are associated

with lower leader contributions. Among participants in the main treatment, the point estimate of the effect of test scores is in fact higher than the effect of period 1 contributions. For two reasons, though, this result is deceptive: first, the range on the contributions variable is larger (from zero to 20) than on the test score variable (from 0 to 15). Second, the variation on the contributions variable is much larger than on the test score variable, as demonstrated in figures B1 and B2. The standardized beta coefficient, which measures how many standard deviations the dependent variable is predicted to change for a one standard deviation change in the independent variable, is .25 for period 1 contributions and -.16 for test score. This indicates that period 1 contributions do in fact explain a larger share of the variation in leader behavior than test score does.

Still, the fact that test score is significant among participants in the main treatment is interesting. The negative coefficient supports the hypothesis that a high test score triggers a feeling of entitlement. This interpretation is further supported by the fact that the estimated effect of test score is also much stronger in treatment 3, where information about period 1 and period 2 actions was revealed, than in treatments 2 and 4, where information was not revealed. In a model controlling for period 1 contributions, the estimated effect of test score is -1.04 in treatment 3, .6 in treatment 2, and -.08 in treatment 4. None of these estimates, however, are significantly different from zero. In sum, leader performance is related to pre-election behavior in a foreseeable way. On the other hand, the share of the variation in leader behavior explained by both measures is not large – the R-squared in the model with both period 1 contribution and test score is only .08. Below, we investigate why this might be the case, and what it means for the effectiveness of political selection.

b) Voting behavior – does voting track predictors of leader behavior?

For political selection to work, it is necessary that voters use the information they are provided effectively. Table 3 investigates how electoral success is correlated with performance in period 1 and 2 for participants in the main treatment. Electoral success is measured by whether or not a participant is elected leader. Since all members of a group have to cast exactly one vote, and all groups have to elect exactly one leader, it is likely that a candidate's performance relative to other group members, rather than absolute performance, determines electoral success. Therefore, the independent variables are both expressed as deviations from group means. The first three models are estimated with OLS (linear probability models), while the last model uses probit estimation which, however, produces very similar results, as do models for the number of votes received (0, 1 or 2) estimated as Poisson- or negative-binomial regressions (not shown). Since all groups must

elect exactly one leader, there is likely to be negative within-group autocorrelation in the residuals. Therefore, standard errors are adjusted for within-group clustering.

Table 3 shows that those who contributed more to the public good in period 1 than others were much more likely to be elected, but those who were smarter (had higher test scores) were not. For example, moving from a minus 10 to a plus 10 point deviation from group mean is estimated to increase the probability of being elected leader by about 60 percentage points. 74 percent of all elected leaders in the core treatment had the highest period 1 contribution in their group (including groups where the leader was tied for the first place with one or two other group members). Only three elected leaders (7 percent) had the lowest period 1 contribution in their group. In contrast, the test score does not predict electoral success. It is nevertheless interesting that the point estimate of the effect of test performance is in all cases negative, mirroring the negative effect of test score on leader contributions presented in Table 2.

Table 3: Determinants of electoral success

	Dependent variable: Elected leader = 1			
	Lin. prob.	Lin. prob.	Lin. prob.	Probit
Period 1 contribution (deviation from period 3 group mean)	0.030 (4.85)***		0.030 (5.02)***	0.034 (3.96)***
Test score (deviation from period 3 group mean)		-0.012 (0.62)	-0.020 (1.12)	-0.024 (1.21)
Constant	0.333	0.333	0.333	
Observations	126	126	126	126
R-squared	0.18	0.00	0.18	

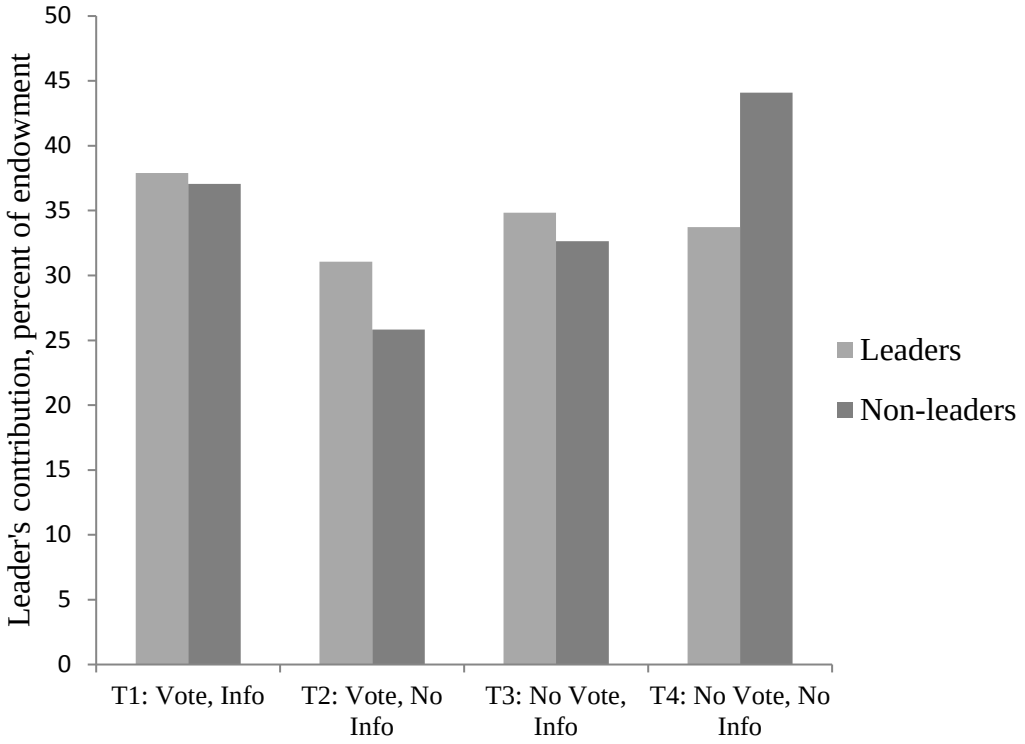
Notes: Absolute value of *t*-statistics in parentheses. Standard errors are adjusted for within-group clustering (for period three groups). Only observations in the main treatment are included. By definition, the constant in these regressions has to be 1/3. The standard error of the constant is therefore zero, and *t*-statistics are not reported. In the probit model, marginal effects are reported. * significant at 10%; ** significant at 5%; *** significant at 1%

In summary, voting behavior tracks the predictors of leader behavior rather well: Both of the variables that voters are informed about before the election have the same sign in the electoral success model as in the leader behavior model, and voters give most weight to the variable which is most strongly related to leader behavior, namely contribution to the public good in period 1.

c) Is political selection effective?

Are the links from period 1 and period 2 behavior to leader performance and voting strong enough to make political selection effective? In other words, do they improve the performance of elected leaders beyond what we would expect if leaders were randomly appointed? The discussion below shows that this is not the case.

Figure 4: Effects of political selection



Note: n = 234. T1 is the main treatment (n = 126).

Figure 4 shows average leader contributions, in percent of the endowment (= 60 points), for elected leaders and others, by treatment. First consider the main treatment (leftmost bars). Elected leaders are not statistically more pro-social than those not elected. The difference between leaders in the main treatment and leaders in the control treatments (who are all effectively randomly appointed) is somewhat larger (4.7 percentage points on average), but also statistically insignificant. In fact, none of the pair-wise differences between two treatments are statistically significant neither

for leaders nor for non-leaders, nor for both groups pooled together.[5] Hence, there is no statistical evidence that political selection improves the performance of elected leaders in the current set-up. Although the information presented to voters is correlated with leader behavior, and voting is in turn correlated with the information variables, these relations are too weak to generate a large effect on the performance of elected decision makers.

We also find no evidence that either voting or providing information about fellow group members in themselves lead to more pro-social behavior. The insignificant effect of information (Leaders in T1&T3 vs. T2&T4, p = .48) is perhaps surprising because information about group members might be expected to trigger feelings of solidarity and identification, and thereby increased leader contributions. For example, Frey and Bohnet (1995) find that dictators give more in dictator games when they are presented with information about the recipient of their allocation. However, our results are in line with Sausgruber (2009) who does not find effects from mutual observation in a public goods game.

Voting might be expected to trigger feelings of reciprocity: In the language of Rabin (1993), I might interpret a vote for me as a "kind" act, which I will respond to with "kindness", in the form of higher contributions. As mentioned in section 2, Corazzini et al. (2014) do find that elected leaders behave more pro-socially than randomly appointed ones.[6] However, there is no evidence that these mechanisms played a role in the present experiment (Leaders in T1&T2 vs. T3&T4, p = .75). The numerous positive contributions to public goods indicate that social preferences play a role but the fact that the *process* by which final outcomes are reached (for example, whether the leader is elected or not) does not seem to make a difference indicates that theories which emphasize agents'

---

[5] Leaders: T1 vs. T2: p = .53, T1 vs. T3: p = .85, T1 vs. T4: p = .60; T1 vs. All controls: p = .52; Non-Leaders: T1 vs. T2: p = .39, T1 vs. T3: p = .62, T1 vs. T4: p = .41, T1 vs. All controls: p = .87; Leaders vs. Non-Leaders in T1 (main treatment): p = .96 (Mann-Whitney tests in all cases).

[6] We might speculate that the use of the strategy method is responsible for the absence of an "effect of democracy". Candidates choose leader strategies before they know whether they are actually elected. At this stage, feelings of obligations towards voters may not be highly salient, because the voting outcome is still hypothetical. However, Corazzini et al. also used the strategy method, and nevertheless found that elected candidates were more benevolent than appointed ones. The key factor behind these divergent results may be that candidates in Corazzini et al. made explicit promises to voters and that the breaking of such promises evokes feelings of guilt. This is the interpretation offered in Corazzini et al., and it is strengthened by the absence of an effect of elections in our experiment, where feelings of guilt were less likely to be prominent, because no explicit promises were made.

preferences over processes, such as theories of reciprocity (e.g. Dufwenberg and Kirchsteiger 2004, Falk and Fischbacher 2006), are less relevant for explaining the observed behavior than theories focusing on agents' preferences over final outcomes, such as theories of inequality-aversion (e.g. Fehr and Schmidt 1999).

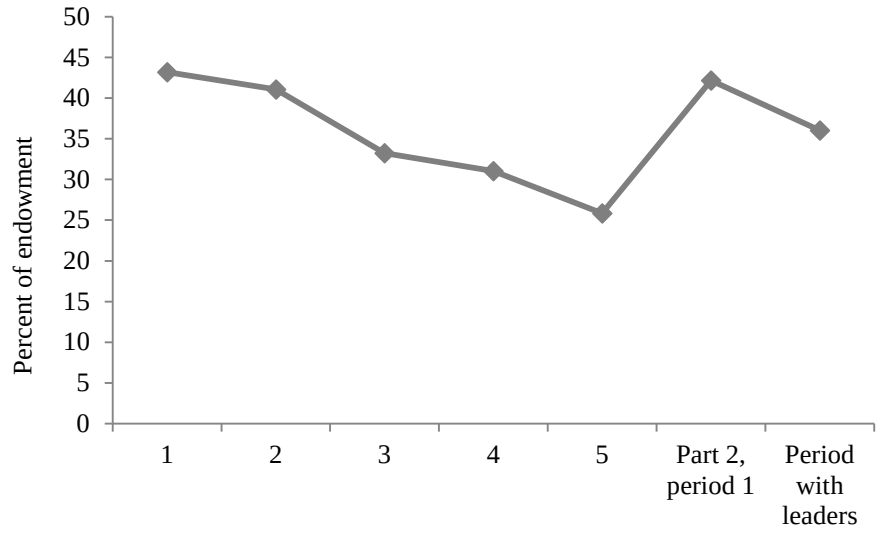
d) Strategic contribution behavior as an explanation for ineffective political selection

The discussion above suggests that strategic contribution behavior by egoists in the pre-election public goods is the reason why the political selection process turned out to be ineffective. In fact, the results presented so far are consistent with the model of political selection and strategic behavior presented in section 4. This theory predicts that there will be a positive, but imperfect correlation between period 1 contributions and leader behavior, and a stronger positive correlation between period 1 behavior and voting, but that political selection might still only have a small influence on the ability of voters to pick pro-social leaders (see Figure 2), due to the effects of strategic contributions by egoists in the pre-election period. These predictions are fully in line with our results. Below, we investigate whether there is more direct evidence of strategic pretending taking place, and if there is, whether it was profitable.

Evidence for strategic contribution behavior

Figure 5 shows average contributions to the public good, in percent of the endowment, for each period in the experiment where participants chose such contributions.

Figure 5: Contributions to the public good



Note: n = 234.

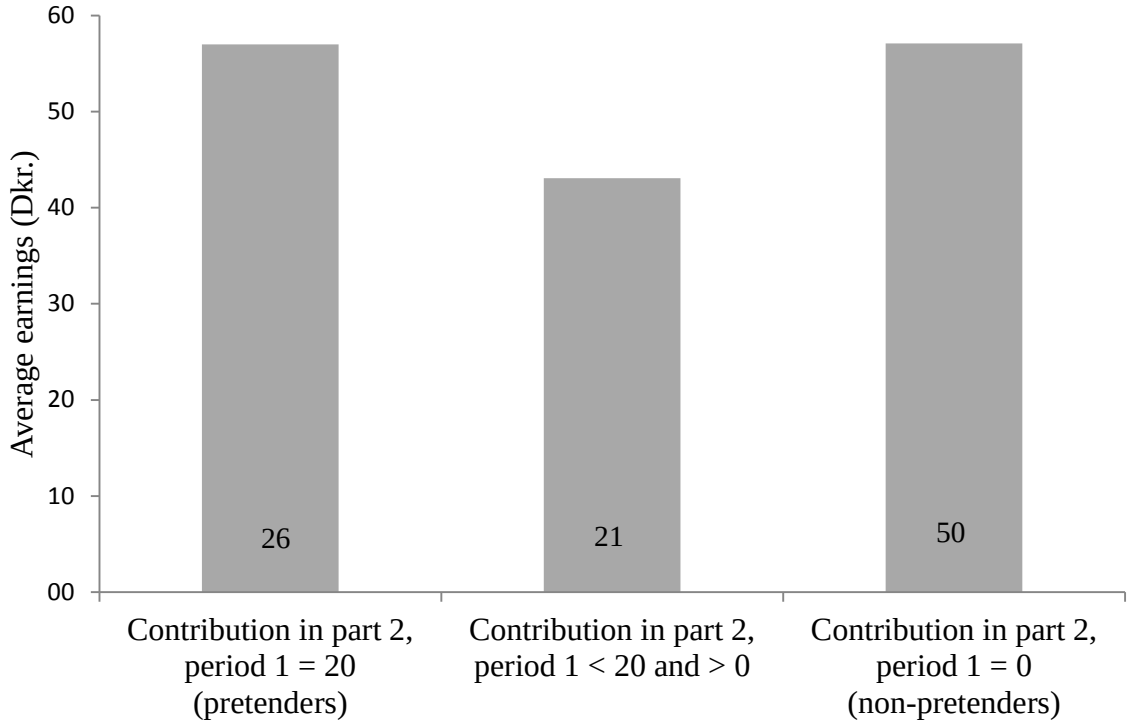
Figure 5 shows that contributions in part 1 exhibit the usual gradual decline (Ledyard 1995), from 43 percent to 26 percent of the endowment. However, in the first period of part 2, where the strategic incentive to contribute is present, contributions jump up to the level of period 1 of part 1. In fact, we find that contributions are significantly higher at the beginning of part 2 than at the end of part 1 ($p < .01$, Wilcoxon signed-rank test). This finding is consistent with the hypothesis that at least some participants increased their contributions for strategic reasons. On the other hand, since there is a break in the experiment between the end of part 1 and the beginning of part 2 for reading instructions, the jump in contributions could also be the result of a "restart effect", which is often observed in experimental sessions interrupted by a break (Cookson 2000). The change in the exchange rate between part 1 and part 2 is another potential factor contributing to differences between the respective periods.

More detailed analysis suggests that strategic motives indeed play a role. We define a "pretender" as someone who contributes 20 points (the maximum amount) in period 1 of part 2, and then contributes nothing as leader. That is, a pretender is an egoist whose behavior in period 1 is indistinguishable from that of a "pro-social type". By this definition, there are 26 pretenders in the experiment, 15 of which were in the main treatment, and 9 of these were elected leaders. Now, if strategic pretending took place, the pattern of shifting from the maximum contribution in one period to the minimum in the next should be more common in period 1 of part 2 and the period with leadership decisions than in any other pair of consecutive periods in the experiment. We find that 11 percent of participants shifted from contributing 20 in period 1, part 2, to contributing zero as leaders. In comparison, less than two percent (four participants) shifted from contributing 20 in period 1 of part 1 to zero in period 2. The differences between the period 1-leadership period pair, and all other pairs of consecutive periods, are all statistically significant ($p < 0.05$, Wilcoxon signed-rank tests).

Did strategic behavior pay?

Figure 6 investigates whether strategic pretending was a profitable strategy for an egoist, i.e. those contributing zero as leaders. The figure shows average, total earnings in the first and last period of part 2 for three groups.

Figure 6: Egoists' earnings in periods 1 and 3 of part 2



Note: Numbers of observations at the bottom of each bar. Only subjects who contributed zero as leaders ("egoists") are included.

The groups are distinguished by how much they contributed to the public good in period 1 of part 2. The first bar includes those contributing 20 points in that period. These are the pretenders. The second bar includes those contributing less than 20 but more than zero, and the third represents those contributing exactly zero. These are the non-pretending egoists. Note that in period 1, the expected earnings of a participant contributing 20 points is 10 points (0.5*20) lower than for participants contributing nothing.

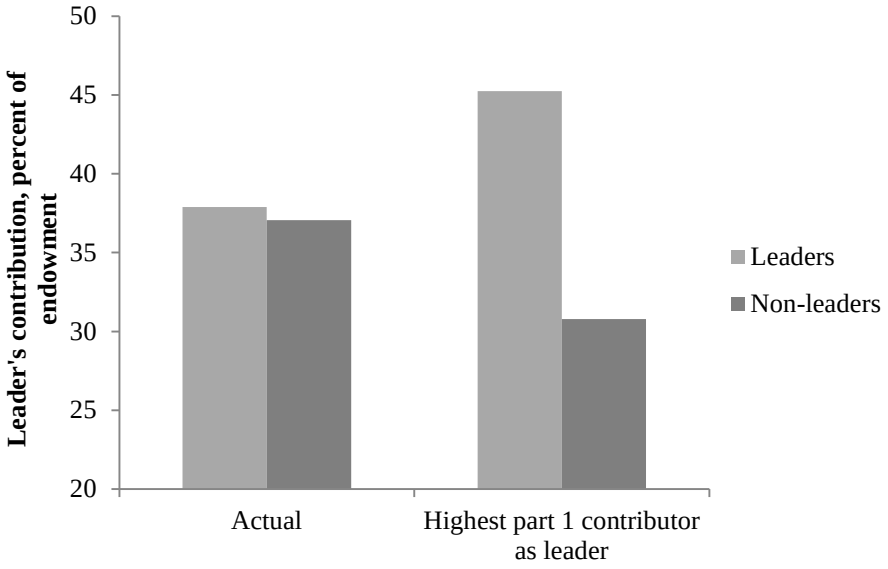
The figure shows that among egoists, total earnings are almost identical for those contributing 0 in period 1 and those contributing 20 (compare leftmost to rightmost bar). High contributions increased the probability of being elected leader and this effect was just strong enough to compensate for the 10 point loss incurred by high contributors in period 1, relative to those contributing nothing. This pattern is consistent with the view that we are observing an equilibrium where some but not all egoists contribute in period 1, and where the share of egoists contributing is just high enough to render the earnings of high- and low contributors equal. This is the theoretical prediction emerging from the model presented in section 4. The only fact not explained by the equilibrium view is that egoists with lower than 20 but higher than zero contributions in period 1

(21 participants) earn significantly less than both those who contribute zero and those who contribute 20. Contributions lower than 20 did not affect election probabilities enough to make up for the loss incurred my making them.

Counterfactual political selection absent pretending

The discussion above has shown that some egoists indeed strategically contribute fully to the public good when they know that they are being observed in the pre-election period (period 1 of part 2), and that pretending to be pro-social was equally profitable as free-riding because voters selected such pretenders sufficiently often. We now ask: how much damage does strategic pretending do to the political selection process? To investigate this question, we construct the counterfactual level of leader contribution that would have resulted were voters not misled by strategic pretending and selected the most pro-social candidate in each group.

Figure 7: Counterfactual contributions (assuming fully effective political selection absent strategic pretending)



Note: Only participants in the main treatment (n = 126) are included.

Figure 7 compares observed to counterfactual contribution levels. The difference between leaders and non-leaders in the counterfactual case is 15 percentage points and highly significant ($p < 0.01$, Mann-Whitney test). Of course, this exercise assumes that all groups would in fact pick

their highest part 1 contributor if they had the chance. Since this might not be the case due to mistakes and strategic voting (see section *d* below), the method may overstate the potential effects of political selection without strategic pretending. On the other hand, the method does not allow groups to exploit information about test scores at all, which potentially leads us to under-estimate the effects of political selection. While this simulation is inferior to an actual, empirical test of the effects of political selection without pretending, it does indicate that strategic pretending plays an important role in rendering the political selection process ineffective.

d) Strategic voting

One potential reason for a less-than-perfect correlation between the candidates' reputation (i.e. period 1 and 2 performance) and electoral success is strategic voting, i.e. voters trying to improve their own chances of being elected by not voting for the candidate they sincerely prefer. Therefore, strategic voting might also be partly responsible for the observed ineffectiveness of the political selection process. The damage from strategic voting results when groups elect leaders whom one would not expect to be the most public-spirited, based on their reputation. We now simulate optimal group behavior by calculating leaders' contribution levels had all groups elected the expected "best" leader. We define "best" leader as the candidate with the highest predicted leader contribution, according to the model in the last column of Table 2. By this criterion, 52 percent of groups in the main treatment elected the optimal leader.

We calculate the difference in contribution levels between leaders and non-leaders, assuming that the best predicted leaders had been elected in all groups in the main treatment and compare it to actually observed levels. We find that counterfactual contributions are eight percentage points higher for elected leaders than for others, but this difference is not statistically significant (p-value = .64, M-W test).

Note that this method provides an upper bound on the effects of strategic voting, since it assumes that *all* failures to elect the predicted best leader are due to strategic voting. This is probably not the case since other behavioral noise (plain errors) may also play some role. In sum, strategic voting may have reduced the effectiveness of the political selection process, but the statistical evidence to support this claim is not strong. Strategic voting seems to have played a smaller role than strategic pretending.

6. Conclusion

This paper investigates how egoistic candidates for leadership positions who pretend to be pro-social can hamper the selection of truly public-spirited (or, pro-social) leaders. We experimentally study a social dilemma situation in which all voters prefer public-spirited leaders, but leaders have incentives to be corrupt, i.e. to use public funds for private gain. In particular, we study a situation in which voters know the candidates' reputation, i.e. how pro-social their behavior was before the election, and candidates are aware of this fact. Hence, egoistic candidates have incentives to act pro-socially when observed, i.e. to pretend to be pro-social, to increase their chances of being elected and to control public funds for private gain. We develop a simple theoretical model of strategic pretending by egoistic candidates for political office and show how pretending, in turn, undermines voters' ability to select public-spirited leaders.

Experimental results show that political selection potentially plays an important role, in the sense that leaders do not all behave alike once in control. While the most common (the modal) behavior is to exploit one's power and to fully appropriate public funds, about 60 percent of leaders show some restraint, and 18 percent of leaders behave in a fully public-spirited manner. Thus, we find substantial variation in leader behavior, and we show that this behavior is partly predictable by variation in reputation, i.e. the amount of pro-social behavior exhibited by candidates before the election. While voters tend to select leaders according to their reputation, the political selection process in the experiment is found to be ineffective. Experimental results are well in line with the predictions of our simple model. We find that an important explanation for the ineffectiveness of political selection is strategic pretending by egoists in the pre-election stage, which severely hampers voters' ability to pick the most public-spirited candidates.

While we think our results are remarkable and clear, we would like to caution the reader from rash extrapolation to more complex settings. In particular, future research should address the following issues.

First, we have chosen the experimental design with the aim to clearly pin down the effects of strategic pretending on political selection. To be able to do so, we needed to investigate a fairly simple setting which abstracts from various factors that may also play a role in political selection in the wild. In a first cut at the issue, as we present it here, it is important to control for these factors since they may interact with strategic pretending in unknown ways. Potentially important factors

that may affect the quality of leader performance that come to mind here are, for example, incumbency, re-appointment incentives and constitutional constraints on the behavior of executives.

Second, our results with respect to the predictive power (and voter preference for leaders who are strong in) cognitive ability vs. social preference should not be extrapolated without further consideration because we think they are context-specific. We find that social preference is the key predictor of whether a leader acts in the group interest, but that cognitive ability is of minor relevance. The reason why this result may be peculiar to our setting is that we study a fairly simple social dilemma situation in which making the "right" choice is essentially a matter of morals and only marginally a matter of cognitive ability. In contrast, one can easily think of trade-offs that arise in the political arena between leaders who are more public-spirited vs. more technically skilled or cognitively able.[7] To address the two issues above, experimenters are called upon to develop richer experimental settings which allow for more dynamics (e.g. incumbency, re-election) and of leadership tasks that require particular skills or are cognitively difficult.

Third, we introduced a number of technical aspects to keep the experimental design simple, and to cleanly identify the effects of strategic pretending. Further experimental research should conduct robustness checks. One possibility is to explore the behavioral implications of our choice to have participants simultaneously play the roles of voters and candidates in the election. The advantage of this procedure is to maximize the number of observations of both voters and candidates. The drawback is that it provides incentives for strategic voting. We found by means of simulation that these incentives have no significant impact on our results, but behavioral robustness checks may nevertheless be comforting. Another possibility is to explore the consequences of our choice to use the strategy method and to study truly sequential decision making instead. We elicit leadership decisions conditional on being elected, i.e. leaders commit to a policy before knowing whether they are actually appointed as leaders. An important advantage of this procedure is that it allows us to compare the behavior of winners and losers in an election. However, it would be interesting to investigate whether leaders behave differently when they know for sure that they have been appointed. For example, one could imagine that the pleasure, and perhaps gratitude,

[7] For example, President Suharto of Indonesia is arguably the most corrupt leader in history (in terms of absolute amounts stolen) but could also boast strong economic growth and an impressive reduction in the Indonesian poverty rate from 60 to 13 percent during his 31 year rule (BBC, March 25, 2004; New York Times, Jan. 28, 2008).

experienced by a leader who knows for sure to have won an election would result in more generous behavior.

References

Acemoglu, D., Egorov, G. and Sonin, K. (2010): Political Selection and Persistence of Bad Governments, *Quarterly Journal of Economics* 125(4): 1511-1575.

Arbak, E. and Villeval, M.-C. (2007): Voluntary Leadership: Selection and Influence, *Social Choice and Welfare* 40(3): 635-662.

Austen-Smith, D. and Banks, J. (1993): Electoral Accountability and Incumbency. In Ordeshook, P.C. (ed.): *Models of Strategic Choice in Politics*. Ann Arbor: The University of Michigan Press.

Banks, J and Sundaram, R. (1993): Adverse Selection and Moral Hazard in a Repeated Elections Model. In W. Barnett, M. Hinich and N. Schofield (eds.): *Political Economy: Institutions, Information, Competition and Representation*. Cambridge: Cambridge University Press.

Barro, R. (1973): The Control of Politicians, *Public Choice* 14: 19-42.

Bernheim, B.D. and Kartik, N. (2014): Candidates, Character, and Corruption, *American Economic Journal: Microeconomics* 6(2) 205-46.

Besley, T. (2005): Political Selection, *Journal of Economic Perspectives* 19(3): 43-60.

Besley, T. (2006): *Principled Agents? The Political Economy of Good Government*. Oxford: Oxford University Press.

Besley, T. and Smart, M. (2007): Fiscal Restraints and Voter Welfare. *Journal of Public Economics* 91(3): 755-73.

Bolle, F., Vogel, C. (2011): Power Comes with Responsibility—Or Does It? *Public Choice* 148:459–470

Brandts, J., Cooper, D.J., Weber, R.A. (2014): Legitimacy, Communication and leadership in the turnaround game, *Management Science* 61(11): 2627-2645.

Brollo, F., Nannicini, T., Perotti, R. and Tabellini, G. (2013): The Political Resource Curse, *American Economic Review* 103(5): 1759–1796.

Caselli, F. and Morelli, M. (2004): Bad Politicians, *Journal of Public Economics* 88: 759-782.

Coate, S. and Morris, S. (1995): On the Form of Transfers to Special Interests. *Journal of Political Economy* 103(6): 1210-35.

Cookson, R. (2000): Framing Effects in Public Goods Experiments, *Experimental Economics* 3: 55-79.

Corazzini, L., Kube, S., Marechal, M.A. and Nicolo, A. (2014): Elections and Deceptions: An Experimental Study on the Behavioral Effects of Democracy, *American Journal of Political Science* 58(3): 579-592.

Dal Bo, E., Dal Bo, P. and Di Tella, R. (2007): Plata o Plomo? Bribe and Punishment in a Theory of Political Influence, *American Political Science Review* 100(1): 41-53.

Dal Bo, P. Foster, A. and Putterman, L. (2010): Institutions and behavior: Experimental Evidence on the Effects of Democracy, *American Economic Review* 100: 2205–2229.

Dasgupta, S. and Williams, K.C. (2002): A Principal-Agent Model of Elections with Novice Incumbents. Some Experimental Results, *Journal of Theoretical Politics* 14(4): 409-438.

Diermeier, D., Keane and M., Merlo, A. (2005): A Political Economy Model of Congressional Careers, *American Economic Review* 95(1): 347-373.

Drazen, A. and Ozbay, E. (2016): Does 'Being Chosen to Lead' Induce Non-Selfish Behavior? Experimental Evidence on Reciprocity. CEPR DP11338.

Dufwenberg, M. and Kirchsteiger, G. (2004): A Theory of Sequential Reciprocity, *Games and Economic Behavior* 47: 268-396.

Falk, A. and Fischbacher, U. (2006): A Theory of Reciprocity, *Games and Economic Behavior* 54: 293-315.

Fearon, J.D. (1999): Electoral Accountability and the Control of Politicians: Selecting Good Types versus Sanctioning Poor Performance. In A. Przeworski, S.C. Stokes and B. Manin (eds.): *Democracy, Accountability and Representation*, Cambridge: Cambridge University Press.

Fehr, E. and Schmidt, K.M. (1999): A Theory of Fairness, Competition and Cooperation, *Quarterly Journal of Economics* 114(3): 817-868.

Ferejohn, J. (1986): Incumbent Performance and Electoral Control, *Public Choice* 50(1): 5-25.

Ferraz, C. and Finan, F. (2009): Motivating Politicians. The Impacts of Monetary Incentives on Quality and Performance. NBER Working Paper 14906.

Fischbacher, U. (2007): z-Tree. Zurich Toolbox for Readymade Economic Experiments, *Experimental Economics* 10(2): 171-178.

Frey, B. and Bohnet, I. (1995): Institutions affect Fairness: Experimental Investigations, *Journal of Institutional and Theoretical Economics* 151(2): 286-303.

Galeotti, F. and Zizzo, D.J. (2015): Competence versus Honesty: What Do Voters Care About?, GATE Working Paper 1520, Lyon Saint-Etienne.

Hamman, J.R., Weber, R. and Woon, J. (2011). An Experimental Investigation of Electoral Delegation and the Provision of Public Goods, *American Journal of Political Science* 55(4): 737-751.

Hobbes, T. (1996 [1651]): *Leviathan*, ed. Tuck, R., Cambridge: Cambridge University Press.

Hoffman, E., McCabe, K., Shachat, K. and Smith, V. (1994) Preferences, Property Rights and Anonymity in Bargaining Games, *Games and Economic Behavior* 7(3): 346-380.

Houser, D. and Stratmann, T. (2008): Selling Favors in the Lab: Experiments on Campaign Finance Reform, *Public Choice* 136(1-2): 215-239.

Houser, D., Ludwig, S. and Stratmann, T. (2009): *Does Deceptive Advertising Reduce Turnout? Theory and Laboratory Evidence*, mimeo, George Mason University.

Houser, D., Morton, R. and Stratmann, T. (2011): Turned Off or Turned Out? Campaign Advertising, Information and Voting, *European Journal of Political* Economy 27(4): 708-727.

Kamei, K., Putterman, L. and Tyran, J.-R. (2015): State or Nature? Formal vs. Informal Sanctioning in the Voluntary Provision of Public Goods, *Experimental Economics* 18(1): 38-65.

Kartik, N. and McAfee, P. (2007): Signaling Character in Electoral Competition, *American Economic Review* 97(3): 852-870.

Kosfeld, M., Okada, A. and Riedl, A. (2009): Institution Formation in Public Goods Games, *American Economic Review* 99(4): 1335-1355.

Ledyard, J.O. (1995): Public Goods Experiments, in Roth, A.E., Kagel, J.H. (eds.): *Handbook of Experimental Economics*, Princeton, NJ: Princeton University Press.

Levy, D.M., K. Padgitt, S.J. Peart, D. Houser and E. Xiao (2011): Leadership, Cheap Talk and Really Cheap Talk, *Journal of Economic Behavior and Organization* 77: 40-52.

Markussen, T. and Tyran, J.-R., (2010): *Serving the Public Interest*, University of Copenhagen, Unpublished Discussion Paper.

Markussen, T., Putterman, L. and Tyran, J.-R. (2014), Self-Organization for Collective Action: An Experimental Study of Voting on Sanction Regimes, *Review of Economic Studies* 81: 301-324.

Markussen, T., Reuben, E. and Tyran, J.-R. (2014), Competition, Cooperation and Collective Choice, *Economic Journal* 124: 163-195.

Matozzi, A. and Merlo, A. (2007): The Transparency of Politics and the Quality of Politicians, *American Economic Review* 97(2): 311-315.

Messner, M. and Polborn, M.K. (2004): Paying Politicians, *Journal of Public Economics* 88: 2423-2445.

Nosenzo, D. and Sefton, M. (2014): Promoting Cooperation: the Distribution of Reward and Punishment Power. In: P. van Lange, B. Rockenbach and T. Yamagishi (eds.): *Reward and Punishment in Social Dilemmas*. New York: Oxford University Press.

Putterman, L., Tyran, J.-R., and Kamei, K. (2011): Public Goods and Voting on Formal Sanction Schemes: An Experiment, *Journal of Public Economics* 95(9-10): 1213-1222.

Rabin, M. (1993): Incorporating Fairness into Game Theory and Economics, *American Economic Review* 83(5): 1281-1302.

Sausgruber, R. (2009): A Note on Peer Effects between Teams. *Experimental Economics* 12: 193-201.

Sutter, M., Haigner, S. and Kocher, M.G. (2010), Choosing the Stick or the Carrot? – Endogenous Institutional Choice in Social Dilemma Situations, *Review of Economic Studies* 77: 1540-1566.

Tyran, J.-R. and L.P. Feld (2006): Achieving Compliance when Sanctions are Non-deterrent, *Scandinavian Journal of Economics* 108(1): 135-156.

Walker, J.M., Gardner, R., Herr, A. and E. Ostrom, E. (2000): Collective Choice in the Commons: Experimental Results on Proposed Allocation Rules and Votes, *Economic Journal* 110: 121-234.